Jet Propulsion Laboratory
California Institute of Technology

Sciences de l'environnement
Institut Pierre Simon Laplace

ESGF CONTAINERS ARCHITECTURE

ESGF F2F WORKSHOP
SAN FRANCISCO (CA)
DECEMBER 2017

LUCA CINQUINI [1]
SEBASTIEN GARDOLL [2]

[1] JET PROPULSION LABORATORY, CALIFORNIA INSTITUTE OF TECHNOLOGY
[2] ENES/IPSL

The ESGF Containers Working Group

- In August 2017, a new "ESGF Containers" working group was formed to provide a unified strategy for evolving the current ESGF architecture into a container-based architecture
- A “container” is a lightweight, standalone package that includes everything needed to run an application (the application, all dependencies, and “just-enough-OS”)
- This working group builds on earlier containerization work supported by the DOE DREAM project, now co-funded by the EU Copernicus project
- Initially targeting a deployment of Docker images via Docker Swarm onto a distributed computing cluster, later evaluating Kubernetes as alternative orchestration engine

Micro-Services Architecture

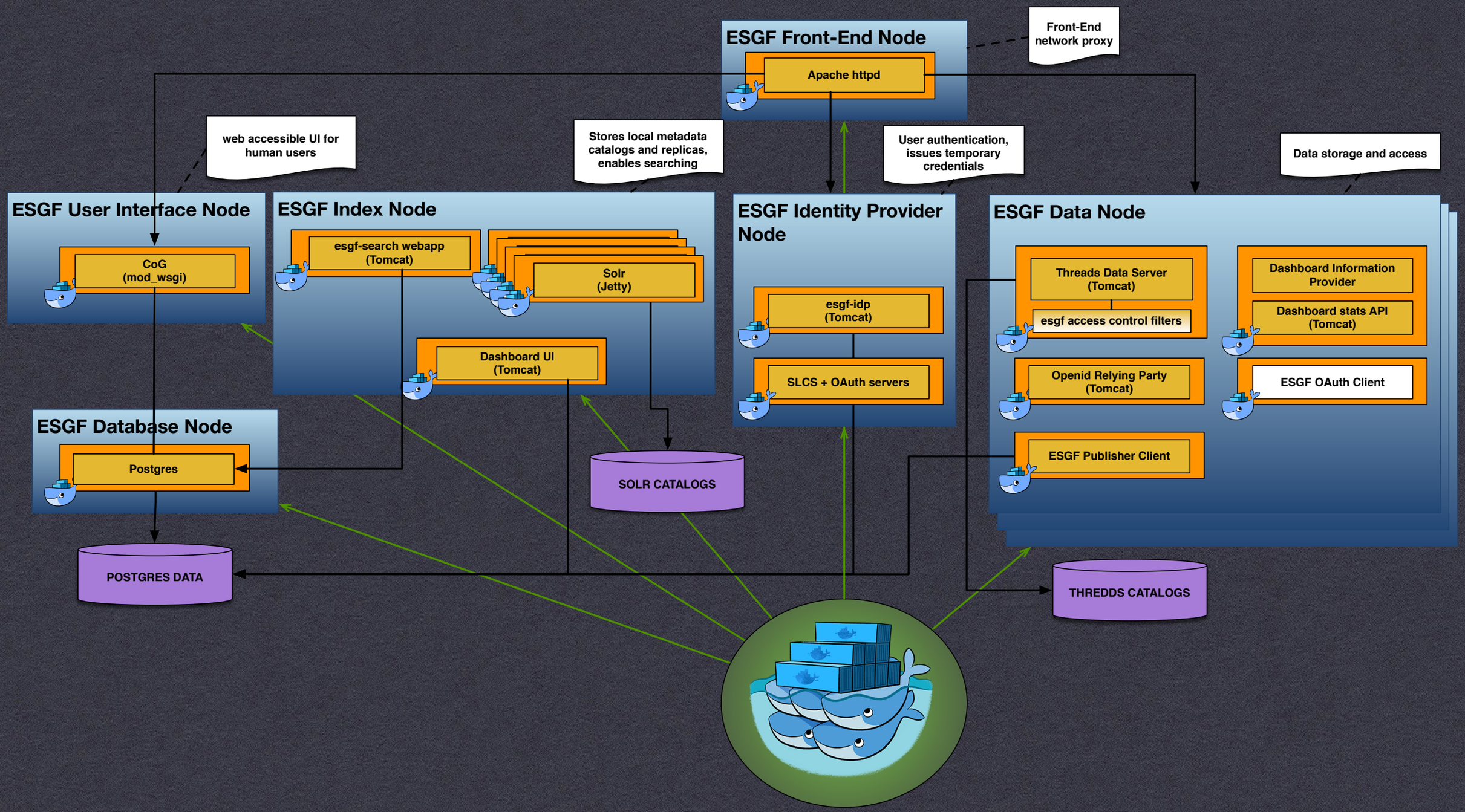
Advantages of container-based architecture ("Micro-Services"):

- Easier to install and upgrade
- Can upgrade separate images
- Can roll back upgrades
- Scalable onto multiple hosts w/ automatic load balancing and failover
- Deployable on laptop, internal cluster or Cloud
- Easier to add new functionality as independent containers
- Flexible deployments of services (as containers) onto distributed hosts

FY17 PROGRESS UPDATE

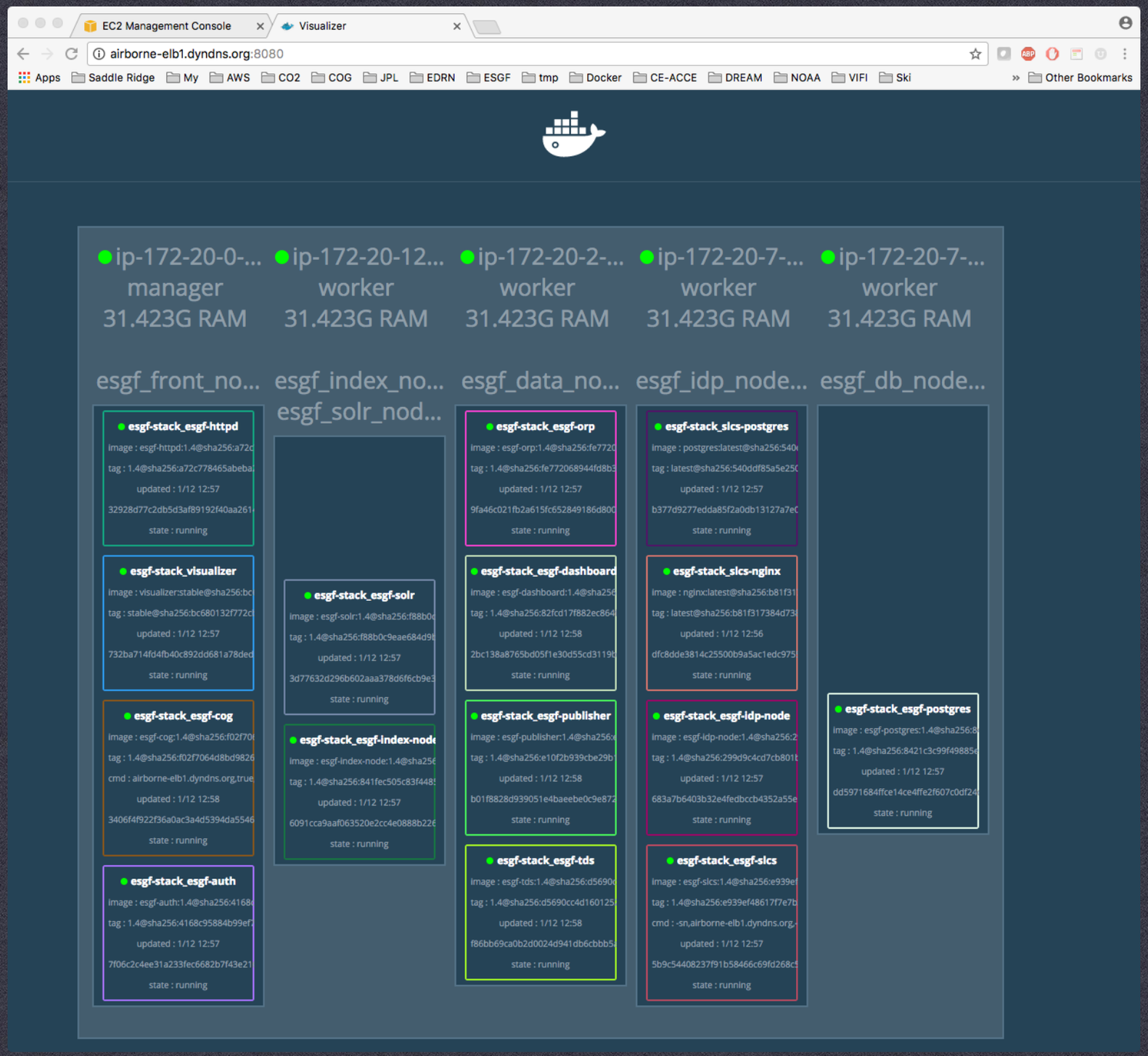
Just released ESGF/Docker version 1.4 which is *almost* a feature complete version of an ESGF Node:

- User registration, authentication and access control
- Data publishing, search and download
- Includes new OAuth2 components: OAuth/SLCS server, ESGF-OAuth client (replacement for ORP), and TDS OAuth filter (disabled by default)
- Includes management of site configuration and sensitive information as Docker configs and secrets
- Not well tested: Node Manager, Dashboard
- Not yet included:
  - GridFTP and Globus Connect Server
  - Live Access Server



ESGF DOCKER ARCHITECTURE V1.4

AS DEPLOYED WITH DOCKER STACK ON 6-NODE SWARM



ESGF DOCKER V1.4 DEPLOYMENT ON AWS-ECS

AS VISUALIZED WITH DOCKER VISUALIZER

FY18 ROADMAP

- Finish integration of current and new ESGF services, including:
  - GridFTP and Globus Connect Server
  - Distributed server-side computing
  - Visus
  - Node Manager
  - Dashboard
  - LAS ?
- Develop a complete testing suite
- Complete transition to OAuth2 authentication
- Deploy ESGF/Docker test nodes at 2+ sites (JPL, IPSL, …) by end of December 2017
- Deploy ESGF/Docker test federation of 3+ sites by Spring 2018
- Transition JPL operations to ESGF/Docker sometimes in 2018
- Evaluate and possibly support deployment with Kubernetes, OpenShift
- If possible, enable more advanced deployment options (more like in 2019…):
  - automatic migration from current ESGF installations to ESGF docker
  - continuous integration (development, testing, distribution, etc.)
  - automatic security updates
  - automatic scalability


7th Annual
Earth System Grid Federation
December 2017
NASA
NOAA
NSF
is-enes
NCI
Face-to-Face Conference
A global consortium of government agencies, educational institutions, and companies dedicated to delivering robust distributed data, computing libraries, applications, and computational platforms for the novel examination of extreme-scale scientific data.

DISCUSSION